Data-point and Feature Selection of Motor Imagery EEG Signals for Neural Classification of Cognitive Tasks in Car-Driving

Anuradha Saha, Amit Konar, Pratyusha Das
Electronics and Telecommunication Engineering
Jadavpur University
Kolkata-700032, India
anuradha.nsec@gmail.com, konaramit@yahoo.co.in, pratyushargj@gmail.com

Basabdatta Sen Bhattacharya
School of Engineering
University of Lincoln
Lincoln, UK
bbhattacharya@lincoln.ac.uk

Atulya K. Nagar
Department of Math and Computer Science
Liverpool Hope University
Liverpool, UK
nagara@hope.ac.uk.

Abstract— **This paper proposes novel algorithms for data-point and feature selection of motor imagery electroencephalographic signals for classifying motor plannings involved in car- driving including braking, acceleration, left steering control and right steering control. Variants of neural network classifiers such as linear support vector machines, and kernel-based support vector machines including radial basis function kernel, polynomial kernel and hyperbolic kernel have been applied to classify the various cognitive tasks. Experimental finding reveals that the proposed data-point and feature selection technique altogether provides better classification accuracies (more than 88%) for all cognitive tasks in comparison with using factor analysis for data-point reduction and feature selection. It is also observed that power spectral density and discrete wavelet transform features are selected among the list of electroencephalographic features for holding the top two rank values for cognitive task classification during car-driving. From the experimental result, it is confirmed that support vector machines with radial basis function along with power spectral density outperforms the remaining feature-classifier pairs in terms of average classification accuracy.**

Keywords— ***data-point selection, feature selection, support vector machine neural network, skewness, differential evolution, motor imagery, electroencephalography.***

I. INTRODUCTION

Car-driving is a complex task involving several motor activities in dynamic environments. Motor activities such as acceleration, braking, steering right control and steering left control during driving includes both motor planning and motor execution. Recent studies [1]-[6] provide a number of methods to classify drivers' cognitive states during driving to avoid traffic fatalities. From the previous literature, it can be confirmed that besides timely motor execution, driver needs to plan correct motor intensions for a particular driving instance, and hence correct classification of motor intensions or planning, i.e. left hand motor imagination during left-hand steering control, right hand motor imagination during right-hand steering control, left leg motor imagination during braking, right hand motor imagination during acceleration from EEG signals are very important concern. This paper, too attempts to classify motor imagery signals for above motor imagination tasks, which in turn gives a clear indication of drivers' cognitive failures during motor planning phase. The present problem is therefore solved by selecting data-point and features of motor imagery EEG signals, and thus minimizing the misclassification rate due to psychological hindrances or cognitive failures.

The first novelty of the present paper is to provide a new approach to data-point reduction. Data point reduction is an important issue in motor imagery-based classification problem in order to select one unique class-representative from a large set of data points (trials). Motor imagery signals, being nondeterministic by nature, does not offer the unique features extracted from several trials of the motor imagery EEG signals captured from the same subject for the similar cognitive task. Therefore, we need to identify the ideal class representative of each data point representing a feature vector of fixed dimension. Previous literature [7, 8] reveal the use of a very popular technique of data-point reduction using PCA. We, here identify one unique representative for each motor imagery class by determining skewness of data-points acquired for that class.

The second novelty, as has been addressed in the paper, is feature selection. An EEG pattern is described by its feature. Feature extraction and selection, therefore are considered as the important steps in motor imagery EEG signal processing. Here, we select well-known EEG feature extraction techniques including time- (Hjorth parameters, Adaptive autoregressive parameters), frequency- (Power spectral density, or in short PSD), and time-frequency (Discrete wavelet transform, or in short DWT) domain techniques to extract motor imagery features from acquired EEG signal depending on the intended

978-1-4799-1959-8/15/$31.00

motor actions. It is important to mention here that classifying motor imagery during driving has always been a serious pattern classification problem, and sometimes, computationally complex feature generates a very high dimensional feature space which may contain both relevant and redundant features. The presence of redundant features adversely affects the performance of the classifier in terms of accuracy and time complexity. Feature selection aims at reducing the dimension of the feature set by retaining only the most relevant features and rejecting the rest. Here, the proposed feature selection technique serves two fundamental purposes.

First, it selects the most useful features from the ideal class-representative, as obtained by each specific feature extraction technique. This is performed by measuring high between-class variance and low within-class variance. It means, for an ideal class-representative obtained after performing each feature extraction technique, only those features which have high discrimination value among different motor tasks will be retained. Here, Differential evolution (DE) is applied to optimally select fewer, i.e. d out of D (where, $d<<D$) number of significant features that correctly classify the various cognitive tasks during driving. Second, optimal number of EEG feature types (such as PSD, DWT and so on) are selected among the different feature extraction techniques for reduced EEG feature sets.

The rest of the paper is organized as follows. In section II, we present preliminaries of differential evolution and skewness. In section III, we provide an overview of the complete system architecture. Section IV offers techniques for the feature selection and feature ranking. Experiments and results are given in section V. Conclusions are listed in section VI.

II. PRELEMINARIES

A. Differential Evolution

Differential Evolution (DE) It is the most widely used evolutionary algorithm due to its inherent merits of low computational overhead, requirement of fewer control parameters and above all its high accuracy. The stages of DE, initialization, mutation, crossover, selection, are discussed in brief.

1) Initialization of the Parameter Vectors

A population of NP D dimensional real-valued parameter vectors is initialised randomly. Each vector of this population is known as *genome/chromosome* which forms a candidate solution to the multidimensional optimization problem. Let the i^{th} vector of the population is $\vec{X}_{i,G} = [x_{1,i,G}, x_{2,i,G}, \ldots\ldots, x_{D,i,G}]$, where G is the generation number and $G = 0, 1\ldots, G\max$.

A uniform distribution is maintained over the total range of $\vec{X}$ during initialisation which helps DE to find the global optima in the search space. So, k-th component of the ith vector can be initialise as

$$x_{k,i,0} = x_{k,\min} + r_{i,k} * (x_{k,\max} - x_{k,\min}) \quad (1)$$

where G=0 for initial population, $r_{i,j}$ is a uniformly distributed random number between 0 and 1,

$x_{k,\max} \in \vec{X}_{\max} = \{x_{1,\max}, x_{2,\max}, \ldots\ldots, x_{D,\max}\}$ and

$x_{k,\min} \in \vec{X}_{\min} = \{x_{1,\min}, x_{2,\min}, \ldots\ldots, x_{D,\min}\}$.

2) Mutation with Difference Vectors

In this stage, three vectors are generated from *parent vector* to the present generation, which are called *target* vectors, a mutant vector obtained through the differential mutation operation is known as *donor* vector and finally an offspring formed by recombining the donor with the target vector is called *trial* vector. The simplest used form of DE-mutation is to create the donor vector for each i-th target vector from the current population. $X_{r_i^1}, X_{r_i^2}$ and $X_{r_i^3}$, three other distinct parameter vectors are selected randomly from the present population. The indices r_i^1, r_i^2 and r_i^3 are randomly chosen different from the range [1, NP], which are also mutually exclusive integers from the base vector index i. One scalar value F is also randomly chosen between (0, 1) for scaling the difference value of the two vectors. Mutation in DE is done by using (2).

$$\vec{X}'_{i,G} = \vec{X}_{r_1^i,G} + F \times (\vec{X}_{r_2^i,G} - \vec{X}_{r_3^i,G}) \quad (2)$$

3) Crossover

After mutation to boost up the diversity of the population, a crossover operation plays an important role after mutation. The donor vector $\vec{X}'_{i,G}$ interchanges its components with the target vector $\vec{X}_{i,G}$ under this operation to form the trial vector $X''_{i,G} = [u_{1,i,G}, u_{2,i,G}, \ldots\ldots, u_{D,i,G}]$. Crossover rate CR is initialised intuitively at the beginning. A random value between (0, 1) is generated, if its less than CR then the donar vector is selected as trial vector else target vector is selected as the trial vector.

4) Selection

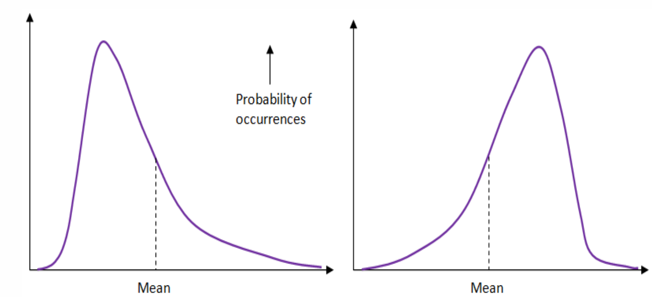
In this stage, DE selects the fittest chromosome to generate the offspring from its parents. Fitness value is calculated using fitness function corresponding to each chromosome of parent population and newly generated trial population. Based on this fitness value, chromosome with maximum fitness value is selected for the new population. Hence, the fitness status of the population either becomes better or remains the same, but never become worsen.

B. Skewness

A *moment* is a precize quantitative measure of the shape of distribution of a set of data points. The third central moment is *skewness* which is a measure of lopsideness of the distribution. In probability theory and statistics, *skewness* is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. Let, there are N number of data points $\{x_1, x_2, ...x_N\}$, mean of these data points is μ and standard deviation is σ, then the skewness for these data points *sk* can be calculated using equation (3).

$$sk = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{x_i - \mu}{\sigma}\right]^3 \quad (3)$$

There are two types of skewness depending on the density of the data points around its mean, *positive skew* and *negetive skew*. It is clear from the Fig. 1 (a) and (b) that if the mass of the distribution is concentrated on the right of the mean i.e. the left tail is longer; then it is *negetive skew* and the *sk* value is also negetive. A positive value of skewness signifies a distribution with an asymmetric tail extending out towards more positive *x*.



(a) (b)

Fig.1. (a) Positive skew, (b) Negative skew

III. SYSTEM OVERVIEW

This section presents basic steps for classifying motor planning during driving from the acquired EEG signals of the subjects (Fig. 2).

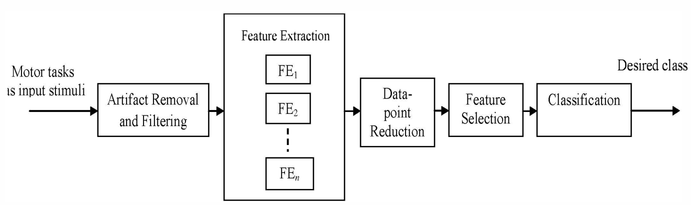


Fig. 2.An overview of motor task classification using EEG signal

Four different kinds of motor planning/imagination during car-driving, including braking, acceleration, steering right, steering left needs to be performed by a subject. The correct planning of motor imagination falls here, as a pattern classification problem and hence the problem is solved by classifying drivers' EEG into correct motor imagination classes. To start with, EEG is first filtered to keep it free from artifacts. The filtered data are sent for feature extraction (FE) to extract the basic primitives of the original signal. Here, FE_1, FE_2, … FE_n represents each specific type of feature extraction techniques. After FE, all data-points are fed to data-point reduction (DPR)/data point selection technique to identify one ideal class representative for reducing computational overhead. After DPR, it has been found that neither all feature types nor all feature values of a specific feature type are equally important for accurate classification of the motor imagery signals. Therefore, we apply a unique feature selection (FS) strategy to select a few significant features among a large set of EEG features. In the second fold of Fs, selection of best two EEG features is held for motor imagery classification during driving. Finally, the selected feature vector for each motor class is applied for classification, which confirms the correct class of cognitive tasks performed.

A. Feature Extraction and Selection

Feature extraction is one of the fundamental step in BCI-EEG. Here, we select four standard feature extraction techniques including PSD, Hjorth parameters, adaptive autoregressive (AAR) parameters and DWT for automatic selection of the most promising feature values as well as type of features that provide best classification accuracy.

B. Data-point Reduction

Data point reduction is an important issue in EEG based classification problem in order to select one unique class-representative from a large set of data points (trials). Here, we too attempt to determine one unique representative for each motor-imagination class by determining skewness for classification of different motor imaginations using different types of EEG features. EEG signals, being nondeterministic by nature, does not offer the unique features extracted from several trials of the EEG signals captured from the same subject for the same stimulus. Therefore, we need to identify the ideal class representative of each data point representing a feature vector of fixed dimension. Let, for each standard driving instance (stimulus, which is used for subject training), a set $A_k = \{\vec{X}_{i,j}, \vec{X}_{i,j}, ..., \vec{X}_{i,j}\}$ is obtained for each type of EEG feature set, where, $i = \{1, 2, \cdots, N\}$ be data-points and $j = \{1, 2, \cdots, D\}$ be the features. Let there be K number of classes, i.e., k lies in [1, K]. Here by determining skewness, we reduce A_k by identifying the representative data point $\vec{\theta}_k$ of dimension $1 \times D$ for each class k =1 to K. The steps for data-point reduction by skewness is given in section III.

C. Classification

Classification of the brain signals is the most vital step in BCI system. In BCI research, the classification algorithms are used to identify the different brain activities based on their signature features. Before applying the BCI system in the real world, the classifier of the system needs to be trained on the prerequisite mental states. Thus, for optimal functioning of the BCI system, the training of the classifier must be optimal. It is difficult to segregate list of standard classifiers into fixed categories, since some classifiers jointly hold one or two characteristics. For example, support vector machines are one type of discriminative classifier, which falls under the broad area of neural network classifiers. In the present paper, we select different variants of SVMs to compare the performance of the classifier using optimally selected features and class-representative.

Linear Support Vector Machines (LSVMs) are particularly suited to handle classification tasks based on drawing separating lines i.e., hyper-planes to distinguish between objects of different class memberships. The use of kernel function along with SVM is important to design non-linear SVM classifier so that the resulting algorithm fits the maximum-margin hyper-plane in a transformed feature space. Since the proposed method is applicable for all hyper-plane based neural network, LSVM and SVM along with three popular kernel functions including Gaussian radial basis function, homogeneous polynomial and hyperbolic tangent are selected for validation the performance of the proposed data-point and feature selection. Eq. (4)-(6) presents the kernel functions of SVM used n this paper.

1. Gaussian radial basis function:

$$k(X_1, X_2) = exp\,(-\gamma \left\| X_1 - X_2 \right\|^2) \text{ for } \gamma > 0 \tag{4}$$

2. Homogeneous polynomial function:

$$k(X_1, X_2) = (X_1 \,.\, X_2)l \text{ , where, } l = number\ of\ polynomials \tag{5}$$

3. Hyperbolic tangent:

$$k(X_1, X_2) = tanh\,(k\,X_1 \,.\, X_2 + c) \in k < 0 \text{ and } c > 0 \tag{6}$$

IV. DATA-POINT AND FEATURE SELECTION ALGORITHM

This section addresses a novel approach of data-point reduction and feature selection, and ranking of features to classify motor imagery signals. Here, the method is two-fold. First is to determine the skewness sk of the hyper-plane-based neural network for selecting one unique class-representative (data-point) from each type of extracted EEG feature set for which the hyper-plane carries wider margin. The skewness of each feature vector for a class is considered as the objective function of Differential evolution (DE) algorithm, where each feature vector represents a chromosome. The optimal solution of DE provides a feature vector, from which d- number of features lying closer to the mean has been selected. Second is to rank all the different feature types and after ranking, feature values of each feature type are updated by multiplying them with the rank of that corresponding feature type.

A. Feature Selection Strategy

1) Feature vector reduction and selection for an individual feature

Let, $\vec{X}_i = \{x_{i,1}, x_{i,2}, \cdots, x_{i,D}\}$ be the n, D- dimensional input feature-vector (data-point), where, i=1 *to* n. The selection of one out of n data-point (ideal class-representative) as well as optimal selection of features of d (d<<D) are presented in Fig. 3. To increase the margin of the hyperplane skewness is determined for all features of both the classes and skew value for $Class_1$ is maximized while skew value for $Class_2$ is minimized using DE algorithm.

This results in reduction of the data points and returns a reduced feature vector (data-point) $\vec{X}_i^N$ of dimension $1 \times D$. Now, d out of D number of features, which are closest to the mean of $\vec{X}_i^N$ are selected from $\vec{X}_i^N$. The selection is based on determining the Euclidean distance $\vec{d}_i$ of feature values corresponding to ideal data point $\vec{X}_i^N$ from its mean.

2) Ranking of individual EEG features

Let there are m number of different types of EEG features. There is a priority vector $\vec{P}$ of length m where $\{p_1, p_2, \ldots p_m\}$ are the data elements of the vector $\vec{P}$. These data elements p_1, p_2, ... p_m are the rank/priority of each EEG feature types. The selected feature values of each feature type are scaled by the rank of that feature type and these scaled features will be fed to the classifier for training and testing. Let the length of the test data set is n_t. Then the objective function, as given in (7) is to be minimized.

$$fit = \sum_{i=1}^{n_t} \left| Actual\ Class(\vec{X}_i) - Obtained\ Class(\vec{X}_i) \right| \tag{7}$$

Ranking of the feature types is confirmed when the difference between fitness values of current trial and the pervious trial has very small value of 0.0001.

Input: Selected feature m number of selected EEG features for all classes (here, $Class_1$ and $Class_2$).
Output: Rank or priority vector of the m EEG feature types.

Step 1: Read values of the control parameters of DE: scale factor F, crossover rate Cr, and the population size NP from user.

Step 2: Set the generation number $t=0$ and randomly initialize a population of NP chromosome $X_t = \{\vec{P}_1(t), \vec{P}_2(t), \ldots\ldots, \vec{P}_{NP}(t)\}$ with $\vec{P}_i(t) = \{p_{i,1}(t), p_{i,2}(t), \ldots.., p_{i,m}(t)\}$ for $i = [1, NP]$ individual uniformly distributed in the range, where $\vec{P}_{\min} = [\vec{P}_{\min-1}, \vec{P}_{\min-2}, \ldots\ldots\ldots, \vec{P}_{\min-m}]$ and $\vec{P}_{\max} = [\vec{P}_{\max-1}, \vec{P}_{\max-2}, \ldots\ldots\ldots, \vec{P}_{\max-m}]$. Evaluate $fit(\vec{P}_i(t))$ for chromosome $\vec{P}_i(t)$, $i = [1, NP]$.

Step 3: WHILE the stopping criterion is not satisfied
DO
FOR i = 1 to NP //do for each individual sequentially
Step 3.1 *Mutation*

Generate a donor vector $\vec{V}_{i,G} = \{v_{1,i,G}, ..., v_{m,i,G}\}$ corresponding to the i^{th} target vector $\vec{P}_{i,G}$ via the differential mutation scheme of DE as:

$$\vec{V}_{i,G} = \vec{P}_{r_1^i,G} + F.(\vec{P}_{r_2^i,G} - \vec{P}_{r_3^i,G})$$

where r_1^i, r_2^i and r_3^i are three random number between 1 to NP satisfying the condition $r_1^i \neq r_2^i \neq r_3^i$

Step 3.2 ***Crossover***

Generate a trial vector $\vec{U}_{i,G} = \{u_{1,i,G}, ..., u_{m,i,G}\}$ for the i^{th} target vector $\vec{P}_{i,G}$ through binomial crossover in the following way:

$$u_{j,i,G} = v_{j,i,G}, \text{ if } (rand_{i,j}[0,1] \leq Cr \text{ or } j = j_{rand})$$

$$= u_{j,i,G}, \text{ otherwise,}$$

Step 3.3 ***Selection***

Evaluate the trial vector $\vec{U}_{i,G}$

Calculate its fitness value *fit* $(\vec{U}_{i,G})$ using (7);

IF $fit(\vec{U}_{i,G}) \leq fit(\vec{P}_{i,G})$

THEN $\vec{P}_{i,G+1} = \vec{U}_{i,G}$

ELSE $\vec{P}_{i,G+1} = \vec{P}_{i,G}$.

END IF

END FOR

Step 3.4 Increase the Generation Count $G = G + 1$

END WHILE

V. Experiments And Results

An emulated driving environment based on virtual-reality scene including a realistic steering wheel, accelerator and brake pedals is used to classify motor intensions required for driving which includes braking, acceleration, steering left and steering right of the participants (Fig. 4). Each driving session takes approximately 14 minutes for each of 12 healthy subjects aged between 22 and 30 and each subject participates in each session for 10 times. All experimental trials are recorded by a stand-alone EEG machine (manufactured by Nihon Kohden) comprising 19 electrodes including Fp_1, Fp_2, F_3, F_4, F_z, F_7, F_8, O_1, O_2, P_3, P_4, P_z, C_3, C_z, C_4, T_3, T_4, T_5 and T_6 and 200Hz sampling frequency. Besides EEG recording, there is a provision to check the possibility of muscle failure using EMG sensors, as has been placed on both hand and leg muscles of the subjects. Fig. 4 presents an experimental set-up for the driving session.

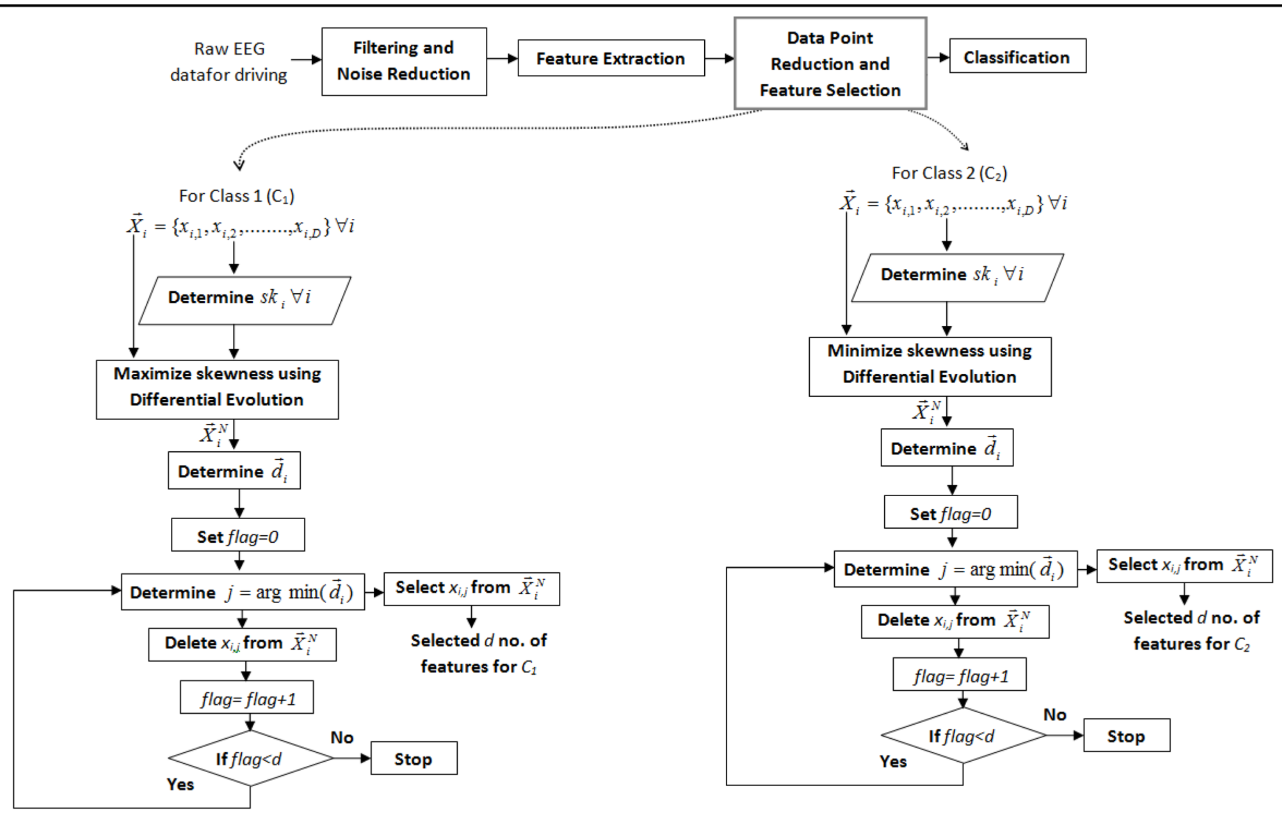


Fig.3. Feature vector reduction and selection for an individual feature

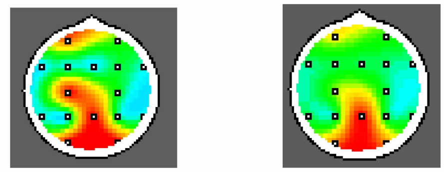
Fig. 4. Experimental set-up for classifying motor planning during driving experiments performed in Artificial Intelligence Laboratory, Jadavpur University.

A. Electrode Selection

To select the important EEG electrodes involved in this driving experiment, we perform the driving session in different driving environments (traffic conditions) (Fig. 5 (a)-(b)) repeatedly and observe the active scalp-maps (Fig. 6 (a)-(b)). Among 19 electrodes (as mentioned above) From Fig. 6 (a) and (b), it is clear that occipital, parietal and motor cortex region possess highest activation during driving instances, thus gaining importance for extracting EEG features from this region of interest. Among this, occipital region is associated with vision. Parietal and motor cortex regions are found significantly active in previous literature [5, 6] during motor planning and/or execution. Therefore, we select P3, P4, Pz, (parietal) and C3, C4 (motor cortex) electrodes for classifying motor intensions.

Fig. 5 (a). Driving condition 1: Sudden increase in traffic

Fig. 5 (b). Driving condition 2: Sudden change in environment (here, heavy rainfall)



(a) (b)

Fig. 6. Active scalp maps for Fig. 5 (a) Driving condition 1 and Fig. 5 (b) Driving condition 2. Here, red color represents the highest activation, whereas blue color represents the lowest activation.

B. Feature Set Preparation using Data-point and Feature Selcetion

This experiment aims at finding the basic primitives (features) from the pre-processed EEG signal. To start with, data matrix of each class represented for determining a specific type of motor operation needs to be prepared. Considering an interval of 10 seconds per subject for a particular driving situation, such as acceleration, braking, steering left, steering right and no action with data samples collected from 12 subjects, each for 15 times, we obtain altogether 180×2000 samples from a specific electrode position for extracting features. Among the acquired data dimension, for each column of the matrix, we obtain 48 PSD EEG features, 276 DWT features, 3 Hjorth parameters and 100 AAR parameters. Therefore, for each motor operation, i.e., for braking, acceleration, steering left or steering right, four sets of distinct type of features have been extracted from each of the significant electrodes. For convenience, plots have been presented by selecting one specific feature (here, PSD) to confirm the feature discrimination between all motor operations.

To give an illustration, Fig. 7(a) and (b) present PSD features extracted from P_3, P_4, C_z C_3, and C_4 electrodes respectively during the corresponding motor tasks. It is already known from the previous literature [5] that P_3, P_4, C_3 and C_4 are associated with the imagined hand movement and C_z is associated with the imagined foot movement, and hence this knowledge has been implemented for classifying motor intensions required for accelerating, braking and steering control. It is apparent from Fig. 7 (a) and (b) that there exist a fewer number of instances where features are jointly capable of discriminating four motor imagery signals. For example, the 12th, 17th, the 23rd, 27th and the 33rd features can be used jointly to classify all the data. From each set of 180 trials for a specific motor imagery task and for each feature type, one ideal class-representative has been obtained. Later, by measuring $\vec{d}_t$, the proposed technique selects 9 PSDs, 22 DWTs, 3 Hjorth parameters and 14 AAR parameters from the extracted feature sets.

C. Confusion Matrix of Average Classification Accuracies using Different Data-point-Feature Selectors

This section uses feature vectors obtained by applying proposed data-point and feature selection algorithm has been applied for comparison with the feature vectors obtained using factor analysis. The confusion matrices of average classification accuracies of different motor imagery signals using the above techniques are presented in Table I. It is clear from Table I that the proposed data-point and feature selection technique altogether provides better classification accuracies (more than 88%) as compared to that of Factor analysis for all motor planning tasks.

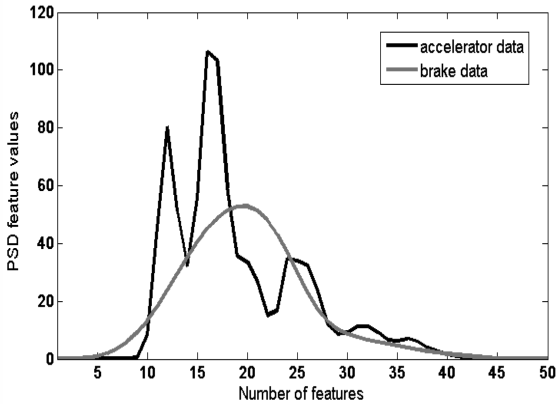


Fig. 7 (a). PSD features extracted from C_z electrode during acceleration and braking respectively

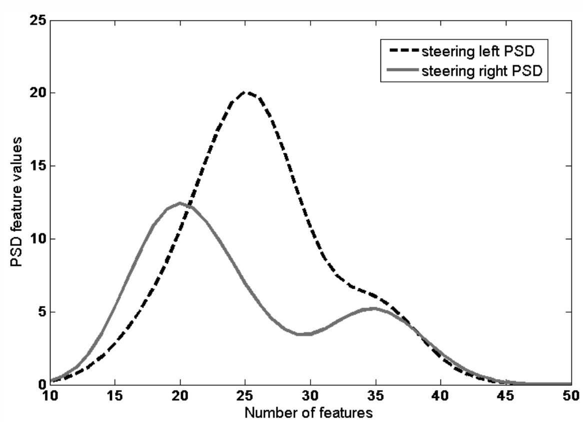


Fig. 7 (b). PSD features extracted from C_3 and C_4 during steering right and steering left respectively

D. Ranking of EEG Features

This section provides the ranking for all EEG feature types for motor imagery classification. After initializing with a random ranking for all feature types, DE automatically updated their ranking as the classification accuracies have been found better for *NP* number of populations offered for all given trials (here, 30). The final rank has been obtained if the fitness function, i.e., classification accuracy of classifiers is optimized. An EEG feature type is selected if it has been found to possess higher ranking in rank vector for optimized classification accuracy obtained from a particular classifier. Table II presents EEG feature type selection through ranking, as the average classification accuracies for different cognitive tasks using LSVM) is given in column 6 of Table II. Table II reveals that for each motor imagery task, PSD and DWT features hold the top two rank values, and hence these two are selected as the most promising EEG features for cognitive task classification during car-driving.

TABLE II
SELECTION OF EEG FEATURES FROM THEIR RANK VALUES TO OBTAIN OPTIMAL CLASSIFICATION ACCURACY

Cognitive Tasks	**Rank of Features**				**Average Classification Accuracy (%)**
	PSD	***DWT***	***Hjorth***	***AAR***	
Braking	0.51	0.37	0.03	0.09	88.00%
Acceleration	0.60	0.33	0.02	0.05	87.53%
Steering left	0.57	0.32	0.04	0.07	90.03%
Steering right	0.62	0.31	0.03	0.04	89.07%

TABLE I
CONFUSION MATRICES OF COGNITIVE TASKS USING THE PROPOSED DATA-POINT AND FEATURE SELECTION TECHNIQUE AND FACTOR ANALYSIS

		Predicted Class using Proposed Technique				**Predicted Class using Factor Analysis**			
Actual Class		***Acceleration***	***Braking***	***Steering left***	***Steering right***	***Acceleration***	***Braking***	***Steering left***	***Steering right***
	Acceleration	92.777	1.111	1.666	2.777	80.777	6.111	4.666	8.777
	Braking	1.111	90.555	4.333	5.000	7.333	82.555	5.333	4.777
	Steering left	2.000	5.333	88.000	4.555	10.333	4.777	78.333	6.777
	Steering right	3.888	5.222	1.222	90.000	8.888	4.222	7.222	80.000

TABLE III
AVERAGE CLASSIFICATION ACCURACIES OF TESTING DATA USING PSD AND DWT FEATURES AND VARIANTS OF NEURAL NETWORK CLASSIFIERS

Cognitive Tasks	**EEG Features**	**Classifiers**			
		LSVM	***SVM-hyperbolic***	***SVM-polynomial***	***SVM-RBF***
Braking	DWT	64.83 (0.041435)	73.15 (0.039964)	84.57 (0.025303)	90.48 (0.012898)
	PSD	60.45 (0.057424)	70.44 (0.031111)	80.36 (0.022429)	**93.05 (0.014978)**
Accelerating	DWT	66.11 (0.049244)	75.00 (0.035146)	89.55 (0.017868)	90.44 (0.013721)
	PSD	62.74 (0.045524)	72.22 (0.07289)	86.33 (0.011079)	**93.11** (0.013281**)**
Steering Left	DWT	71.77 (0.055172)	73.44 (0.042126)	84.77 (0.033340)	90.55 (0.011700)
	PSD	65.44 (0.068212)	76.22 (0.05946)	80.55 (0.018120)	**92.77 (0.012549)**
Steering Right	DWT	72.22 (0.056788)	78.00 (0.041941)	86.55 (0.036689)	90.33 (0.013261)
	PSD	71.22 (0.062178)	77.33 (0.055521)	83.11 (0.042959)	**93.55 (0.012931)**

E. Classifier Performance

This section provides the classification accuracies of four standard neural network classifier including LSVM, SVM-RBF, SVM-polynomial and SVM-hyperbolic after data-point and feature selection. The classification accuracies are obtained after performing 10-fold classification, where, 9 out of 10 fold is used for training and the remaining fold is applied for validation. Table III lists the average classification accuracies for all four classifiers during testing, where the highest average classification accuracy for PSD and SVM-RBF as the best feature-classifier pair is marked in bold.

VI. CONCLUSIONS

This paper concludes its three important findings: 1) the proposed data-point and feature selection technique altogether provides better classification accuracies (more than 88%) for all cognitive tasks in comparison with using factor analysis for data-point reduction and feature selection, 2) the proposed technique performs optimal selection of two out of four EEG features, such as PSD and DWT for cognitive task classification during car-driving, 3) SVM-RBF classifier outperforms the remaining classifiers in terms of average classification accuracy (91.78%), whereas PSD and SVM-RBF has been proved to be the best feature-classifier pair in terms of average classification accuracy.

VII. ACKNOWLEDGMENT

Funding by University Grant Commission, India for the Project: University with Potential for Excellence Program in Cognitive Science (Phase II) granted to Jadavpur University, India is gratefully acknowledged.

REFERENCES

[1] F. C. Lin, L. W. Ko, C. H. Chuang, T. P. Su, and C. T. Lin, "Generalized EEG-based Drowsiness Prediction System by Using a Self-Organizing Neural Fuzzy System," *IEEE Transactions on Circuits and Systems I-REGULAR PARegular Papers*, vol. 59, no. 9, pp. 2044-2055, 2012.

[2] S. Kar and A. Routray, "Effect of Sleep Deprivation on Functional Connectivity of EEG Channels," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol.43, no.3, pp.666-672, May 2013.

[3] Q. Ji, P. Lan, and C. Looney, "A probabilistic framework for modeling and real-time monitoring human fatigue," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 36, no. 5, pp. 862-875, 2006.

[4] C.T. Lin, L. W. Ko, & T. Shen, "Computational Intelligent Brain Computer Interaction and Its Application," *IEEE Computational Intelligence Magazine*, November 2009**.**

[5] A. Saha, A. Konar, R. Burman and A. Nagar, "EEG Analysis for Cognitive Failure Detection in Driving Using Neuro-Evolutionary Synergism," *in theProceedings of IEEE International Joint Conference on Neural Networks (IJCNN)*, pp. 2108-2115, Beijing, China, July 7-11, 2014.

[6] A. Saha, S. B. Roy, A. Konar, R. Jonarthanan, "An EEG-based Cognitive Failure Detection in Driving Using Two-stage Motor Intension Classifier", *in the Proceedings of IEEE International Conference on Control, Instrumentation, Energy & Communication (CIEC)*, pp. 277-281, Kolkata, January 31-February 2, 2014.

[7] A. Saha, A. Konar, P. Rakshit, A. L. Ralescu and A. K. Nagar, "Olfaction recognition by EEG analysis using differential evolution induced Hopfield neural net", *in the Proceedings of IEEE International Joint Conference on Neural Networks (IJCNN)*, pp. 1-8, Dallas, August 4-9, 2013.

[8] A. Saha, A. Konar, A. Ralescu and A. K. Nagar, "EEG analysis for classification of Olfactory signals Using a Recurent Neural Classifier," *IEEE Trans. on Human-Machine Systems*, vol. 44, no. 6, pp. 717-730, December, 2014.